2K

УДК 519.651.2

А.А. Каленчук-Порханова
Институт кибернетики им. В.М. Глушкова НАН Украины, г. Киев
ioanna@public.icyb.kiev.ua

Наилучшая чебышёвская аппроксимация как эффективный способ обработки информации

Рассматривается проблема наилучшей чебышевской аппроксимации для эффективной обработки информации. Дается обоснование преимуществ разработанных алгоритмов аппроксимации, которые связаны с их оптимизацией по точности и быстродействию.

Введение

Решение проблемы информатизационного обеспечения общества является всё более актуальным и определяющим условием его эффективного развития. Информация о состояниях исследуемых процессов в природе и обществе как результат проведения измерений является, как правило, *дискретным представлением в виде информационных массивов числовых данных* о функциональных зависимостях, характеризующих эти процессы. Трудности, а зачастую и невозможность использовать массивы возникают при решении таких прикладных задач, как математическое моделирование и прогнозирование, определение функциональной закономерности природы исследуемого процесса (например, в виде эмпирической формулы), и при решении проблем экономного хранения больших по объему массивов данных либо скоростной передаче по каналам связи, а также при возобновлении значений функциональной зависимости на «*неосвещённых замерами*» участках (*обратная задача*) и др. Эти трудности преодолеваются путём замены (*сжатия*) больших массивов аналитическими выражениями с небольшим количеством параметров (*прямая задача*).

Для решения этой проблемы применяются средства *аналитической обработки данных* на основе методов и алгоритмов аппроксимации (приближения) функций. Предлагается эффективный способ аналитической обработки данных, обеспечивающий высокую точность приближения и использующий аппроксимирующие аналитические выражения разных типов [1]. Этот способ основан на методах и алгоритмах *чебышёвской наилучшей аппроксимации*, который имеет существенное преимущество перед интерполяционным и среднеквадратическим способами приближения. Причём главное преимущество этого способа – обеспечение требуемой гарантированной точности приближения во всей непрерывной области задания функции позволяет сохранять полученную точность во всех точках при решении обратной задачи.

Разработанные автором алгоритмы основаны на методе Е.Я. Ремеза последовательных чебышёвских интерполяций (п.ч.и.) [2] и, в свою очередь, обладают преимуществами перед алгоритмами других авторов, так как для них проведен анализ всех видов погрешностей с целью их *оптимизации по точности и быстродействию* [3], [4].

Основные задачи аналитической обработки данных. В зависимости от цели аналитической обработки массивов числовых данных можно выделить такие основные задачи, для решения которых используются методы и алгоритмы аппроксимации функ-

ций. В то же время следует заметить, что с точки зрения теории аппроксимации эти задачи тесно связаны между собой и решение одной из них часто используется для решения других.

Задача аналитичного представления массивов числовых данных состоит в приближении дискретно заданной функциональной зависимости f различными типами аналитических выражений (аппроксимантов) F с целью их эффективного использования при решении различных прикладных проблем.

Задача поиска эмпирических закономерностей процесса, выраженного экспериментальными данными, с целью построения эмпирических формул состоит в том, что полученная таблица значений измеряемых величин x_i и y_i $(i=\overline{1,N})$ является приближенным представлением некоторого эмпирического закона $F(x;A)$, который характеризуется небольшим количеством параметров $A=\{a_0, a_1,..., a_n\}$ и для определения которого используется тот или иной способ приближения.

Задача сжатия численной информации в виде больших по объему массивов числовых данных состоит в приближенной замене дискретно заданной функции f на одномерной или многомерной сетке некоторым аналитическим выражением F с небольшим числом параметров. Эта задача характеризуется коэффициентом сжатия C, который определяется формулой

$$C = b(f)/b(F),$$

где $b(f)$, $b(F)$ – число бит, необходимых для сохранения соответственно дискретно заданной функции f и аппроксиманта F.

Задача приближенного восстановления значений дискретно заданной функции на «неосвещенных замерами» участках (обратная задача) возникает, как правило, когда значения функции f получены только в некоторых точках $x_1, x_2, \ldots, x_N$ из области определения S, и которых оказывается недостаточно для проведения дальнейших исследований. В таком случае применяется способ приближенного представления функции f аналитическим выражением с целью его использования для вычисления недостающих значений функции в других точках из S. На практике необходимость решения этой задачи возникает, когда при исследовании соответствующих процессов и объектов получение дополнительных данных затруднено или вообще невозможно, например, при исследовании природных процессов, а также при переходе от нерегулярной сетки к регулярной для картографо-математического моделирования в географических информационных системах (ГИС'ах).

Задача сглаживания экспериментальных данных возникает в случаях, когда функция f задана на сетке S своими значениями, которые могут иметь значительные случайные погрешности измерений. В случае, когда предварительно известны свойства поведения исходной функции, используется способ сглаживания экспериментальных данных с целью устранения указанных погрешностей.

Чебышевский способ аппроксимации для аналитической обработки данных. Наиболее распространенными на практике для аналитической обработки данных являются методы и алгоритмы приближения функций, которые основываются на различных способах аппроксимации, а именно: интерполяционном, среднеквадратичном и равномерно-чебышевском.

По сравнению с интерполяционным и среднеквадратичным способами приближения *способ наилучшей чебышевской аппроксимации* является *наиболее эффективным и универсальным* и обладает особенным свойством не только получать высокую

точность аппроксимации в точках дискретного представления функциональных зависимостей, но и обеспечивать *требуемую гарантированную точность приближения* во всех точках непрерывного интервала их задания [4].

Проблема наилучшей (равномерной) чебышевской аппроксимации функции $f(x)$ на интервале $[a,b]$ основана на чебышевском принципе минимизации величины меры равномерного приближения $L[H_n] \equiv \max_{x\in[a,b]} |f(x) - H_n(x;A)|$ и состоит в нахождении такого аппроксиманта степени n с набором коэффициентов $A = (a_0, a_1, \ldots, a_n)$ из всей совокупности аппроксимантов H_n степени $\leq n$, который удовлетворяет условию минимакса $L[H_n] \equiv \min$, где $f(x)$ – непрерывная на $[a,b]$ функция и $\min_{H_n} L[H_n]$ – наименьшее возможное значение меры равномерного приближения.

В качестве $H_n(x;A)$ будем рассматривать классы P_n всех полиномов степени не выше n вида $P_n(x) = \sum_{i=0}^{n} a_i x^i \equiv P_n(x;A)$ и классы r_n всех дробно-рациональных выражений порядка $n = l+m$ вида $r_n(x) = P_l(x)/Q_m(x) \equiv r_n(x;A;B)$, где $P_l(x)$ и $Q_m(x)$ – полиномы степеней l и m с наборами соответственно коэффициентов $A = \{a_i\}, i = \overline{0,l}$ и $B = \{b_j\}, j = \overline{0,m}$. Тогда постановка задачи нахождения наилучших чебышевских аппроксимантов имеет следующий вид:

$$\max_{x\in[a,b]} |f(x) - P_n(x;A)| \equiv L[P_n] = \min_{P_n}, \qquad \min_{P_n} L[P_n] = L[\Pi_n(x)] = \rho, \tag{1}$$

$$\max_{x\in[a,b]} |f(x) - r_n(x;A;B)| \equiv L[r_n] = \min_{r_n}, \qquad \min_{r_n} L[r_n] = L[R_n(x)] = \delta, \tag{2}$$

где $\Pi_n(x)$ и $R_n(x)$ – полиномиальный и дробно-рациональный наилучшие чебышевские аппроксиманты, а ρ и δ – соответственно величины их наилучших приближений.

Существование, единственность и свойства наилучших аппроксимантов вытекают из классических теорем Э. Бореля и П.Л. Чебышева – для полиномиальных и Н.И. Ахиезера и П.Л. Чебышева – для дробно-рациональных аппроксимантов [5]. На основании этих теорем единственные решения задач (1) и (2) совпадают соответственно с решениями «элементарных» задач вида

$$\max_{x\in X_1} |f(x) - P_n(x;A)| = \dot{\rho}, \qquad \max_{x\in X_2} |f(x) - r_n(x;A;B)| = \dot{\delta} \qquad (1'), (2')$$

на таких (n+2)-точечных подмножествах $X_1, X_2 \subset [a,b]$, для которых величины $\dot{\rho}$ и $\dot{\delta}$ достигают свои наибольшие возможные значения, равные ρ и δ.

Каждая из таких (n+2)-точечных задач называется *задачей чебышевской интерполяции* функции $f(x)$ на множестве (n+2)-х точек, которые являются соответственно *чебышевским альтернансом* для полиномиальной задачи (1) и *экстремальным базисом* для дробно-рациональной задачи (2). Именно это «замечательное» свойство «чебышевского альтернанса» является теоретической основой для нахождения всех наилучших чебышевских приближений.

Все известные способы решения чебышевской задачи можно в основном разделить на способы, основанные на распространении методов Е.Я. Ремеза (особенно 2-го), и способы, использующие аппараты линейного и выпуклого программирования. Для

решения дробно-рациональной задачи используются также различные приемы последовательной дифференциальной линеаризации по параметрам-коэффициентам. Синтезом многих направлений в последние годы явилась также теория сплайновой аппроксимации [5].

Преимуществами методов Е.Я. Ремеза являются сравнительно быстрая скорость их сходимости (в некоторых случаях квадратичная) и возможность стандартизации вычислений, что очень важно для эффективности их численных реализаций.

Метод Е.Я. Ремеза решения задач (1) и (2) основан на последовательных чебышевских интерполяциях (п.ч.и.), r шагов которых сводятся к нахождению последовательности (n+2)-точечных S-наборов $S_r = \{x_\nu^{(r)}\}, \nu = \overline{0, n+1}$, сходящейся к искомому чебышевскому альтернансу или экстремальному базису, и решению на каждом j-м шаге систем алгебраических уравнений

$$f\left(x_\nu^{(j)}\right) - P_{n,j}\left(x_\nu^{(j)}\right) = (-1)^\nu \cdot \rho'_j, \tag{3}$$

$$w\left(x_\nu^{(j)}\right) \cdot \left[f\left(x_\nu^{(j)}\right) - P_{l,j}\left(x_\nu^{(j)}\right) / Q_{m,j}\left(x_\nu^{(j)}\right)\right] = (-1)^\nu \delta'_j \tag{4}$$

соответственно линейных относительно коэффициентов $a_k, k = \overline{0, n}$ полинома $P_{n,j}(x)$ и величины ρ'_j в задаче (3) и нелинейных относительно коэффициентов $a_i, i = \overline{0, l}$ и $b_i, i = \overline{0, m}$ и величины δ'_j в задаче (4), $x \in S_j$, $\nu = \overline{0, n+1}$.

Следует заметить, что в отличие от полиномиального случая сходимость чебышевских интерполяций в случае задачи (4) теоретически обеспечивается не при любом начальном наборе (n+2)-х точек, хотя практика применения известных в литературе различных численных реализаций даже упрощенных вариантов метода Ремеза показала крайне редкую их «несходимость».

Основная трудность всех численных реализаций п.ч.и. состоит в выборе (n+2)-точечных подмножеств области аппроксимации, на которых осуществляются шаги чебышевских интерполяций. От способа этого выбора зависит не только скорость сходимости всего метода, но и сам факт его сходимости. Возможны три класса вариантов последовательной замены указанных (n+2)-точечных наборов: *оптимальный*, *полуоптимальный* и *допустимый* варианты.

При оптимальном варианте для некоторых классов функций обеспечивается квадратичная скорость сходимости, что на практике обеспечивает, как правило, всего одну-две итерации. В *полуоптимальном* варианте – число итераций для получения аналогичного эффекта оказывается в несколько раз больше, а в *допустимом* оно может быть во много раз большим [2, с. 79].

Алгоритмы реализации способа наилучшей чебышевской аппроксимации. Разработанные в ИК им. В.М. Глушкова НАНУ алгоритмы основаны на втором методе п.ч.и. Е.Я. Ремеза и обладают рядом преимуществ по сравнению с известными в литературе аналогичными численными реализациями [6]. Главным достоинством этих алгоритмов является то, что в них разработан *оптимальный* способ замены (n+2)-точечных наборов при переходе к новому S набору [7].

Численная реализация разработанных алгоритмов также имеет ряд дополнительных преимуществ, связанных с оптимизацией этих алгоритмов по точности и быстродействию [8]. Алгоритмы могут находить либо аппроксимант заданной фиксированной степени (вход по степени), либо такой аппроксимант, который обеспечивает

заданную точность приближений (вход по точности), которую не должна превышать апостериорная оценка полной погрешности аппроксимации по соответствующему алгоритму. При этом при исследовании поведения функции отклонения $w(x)(f(x)-F(x;A))$ во всех точках множества S на каждом шаге чебышевских интерполяций учитывается как *верхняя*, так и *нижняя* границы величины наилучшего приближения, что позволяет получить более точную оценку полной погрешности алгоритмов.

Для решения полиномиальной задачи разработаны два алгоритма («А» и «Б»), соответствующие записям аппроксимирующего полинома в формах $\sum_{i=0}^{n} a_i\, x^i$ для алгоритма «А» и $\sum_{i=0}^{n} b_i\; T_i(x)$ для алгоритма «Б».

Алгоритм «Б» возник как модификация алгоритма «А» в связи с работой Н.С. Бахвалова [9] и может быть использован как существенное улучшение алгоритма «А», так как используемая в алгоритме «А» привычная запись полинома в виде $\sum_{i=0}^{n} a_i x^i$ в случаях большого «разброса» значений коэффициентов a_i может стать источником большой погрешности округлений при вычислении значений полинома в точках по схеме Горнера. Предлагаемый Н.С. Бахваловым алгоритм использует запись многочленов в виде линейной комбинации многочленов Чебышева, а именно $Q_n(x)=\sum_{i=0}^{n} b_i T_i(x)$, и позволяет существенно уменьшить указанную погрешность при вычислений значений полинома.

В результате анализа полных погрешностей алгоритмов «А» и «Б» установлено, что преимущество алгоритма «Б» становится ощутимым при степенях $n>0$; при $n\le 0$ оба эти алгоритма примерно равносильны.

Алгоритмы могут работать как для случая аналитически заданной, так и для случая дискретно заданной функции $f(x)$. В обоих случаях процедура выбора (n+2)-точечного набора на каждом шаге ч.и. предполагает дискретизацию, поэтому отличие этих случаев только в том, что во втором из них значения функции в нужных точках известны, а в первом предусматривается процедура их вычисления.

Дробно-рациональную аппроксимацию целесообразно использовать для повышения точности приближения функции $f(x)$, природа которой такова, что она на некоторых участках имеет «всплески». В этом случае удается более точно учитывать особенности поведения функции.

В силу того, что переход от полиномов к дробно-рациональным выражениям расширяет класс чебышевских аппроксимантов, очевидно, что $\delta_n \le \rho_n$, где δ_n и ρ_n – соответственно величины наилучшего дробно-рационального и полиномиального приближений той же функции на том же самом отрезке аппроксимации E .

Несмотря на то, что переход от полиномиальных аппроксимантов к дробно-рациональным не позволяет увеличить порядок стремления к нулю величины δ_n для всего класса функций, можно утверждать, что благодаря этому переходу расширяется класс функций, для которых величина наилучшего приближения будет того же порядка, что и для «хороших» функций.

Следует заметить, что в отличие от полиномиального случая сходимость п.ч.и. (причем, как правило, квадратическая) для дробно-рациональной задачи согласно работам А. Ральстона [5] может быть обеспечена только при условии наличия начального приближения $R_n^{(0)}(x)$, «близкого» к наилучшему дробно-рациональному аппроксиманту $R_n(x)$.

В некоторых частных случаях для многих математических функций, в особенности при малых значениях $n = m + k$, в качестве начального набора точек для нахождения $R_n^{(0)}(x)$ можно применять точки уклонения полинома Чебышева, но для общего случая чебышевской рациональной задачи нет никаких практических рекомендаций относительно выбора «близкого» $R_n^{(0)}(x)$. Этот недостаток устранен работами Вернера [5], который предложил метод, всегда сходящийся с любого начального приближения. Однако громоздкость вычислений и низкая скорость сходимости (линейная), по утверждению самого автора, не позволяют использовать его эффективно на практике.

Учитывая сказанное выше, а также то, что даже упрощенные варианты п.ч.и. зачастую сходятся, автором был разработан сочетающий в себе преимущества методов Ремеза и Вернера *комбинированный алгоритм* (к.а.) [5]. Идея этого алгоритма состоит в том, что на практике часто желательно получить приближение, модуль-максимум уклонения которого не превышает некоторое заданное число $\gamma > 0$. Поэтому естественно вычислять последовательность аппроксимантов, начиная с более малых значений степеней l и m и повышая эти значения до получения приближения с желаемой точностью. Таким образом, будем предполагать, что всегда перед началом вычисления $R_{l,m}(x)$ уже получен наилучший аппроксимант $R_{l-1,m-1}(x)$. Для получения аппроксимантов каждой новой степени вначале используется метод п.ч.и. с обязательной проверкой сходимости алгоритма на каждом шаге ч.и. В случаях, когда сходимость не нарушается, метод п.ч.и. работает до конца, т.е. до получения наилучшего аппроксиманта текущей степени. Если же на каком-то шаге п.ч.и. сходимость нарушается (например, число участков перемены знака уклонений меньше $l + m + 2$), то для получения начального приближения $R_{l,m}^{(0)}(x)$, имеющего не меньше $l + m + 2$ экстремума, вступает в работу *начальный алгоритм* (н.а.) метода Вернера. После этого снова работает метод п.ч.и. Если не учитывать случаи вырождения, сходимость метода п.ч.и. обеспечена и абсолютное значение максимального уклонения полученного наилучшего аппроксиманта не превышает величину желаемой точности, то алгоритм заканчивает свою работу. В противном случае степень аппроксиманта повышается и снова начинает работать метод п.ч.и.

В разработанном к.а. учитываются случаи *вырождения*, когда у наилучшего аппроксиманта предыдущей степени уклонений равной величины и чередующих знак больше $l + m + 2$, и *почти вырождения*, когда участков перемены знака больше $l + m + 2$, а модуль-максимумов уклонений равной величины точно $l + m + 2$.

Несмотря на то, что случаи вырождения и особенно почти вырождения крайне редки, очень важно с вычислительной точки зрения уметь их распознавать до начала вычисления вырожденного аппроксиманта. С этой целью в алгоритме предполагается, что перед вычислением $R_{l,m}(x)$ всегда имеется наилучший аппроксимант $R_{l-1,m-1}(x)$. Это равносильно тому, что вычисления следует начинать всегда с невырожденных аппроксимантов $R_{l-m,0}(x)$ для $l \geq m$, или $R_{0,m-l}(x)$ для $l < m$.

В случае дробно-рациональной аппроксимации исходная функция $f(x)$ предполагается заданной на $[a,b]$ как аналитически, так и дискретно, однако при поиске новых экстремумов уклонений используется дискретная процедура вычисления значений $f(x)$ в точках. В отличие от полиномиального случая, когда вход в алгоритм может начинаться либо по заданной точности, либо по степени, вход в к.а. осуществляется *только по заданной точности* γ. Отличие также состоит в том, что в этом случае система $l+m+2$ алгебраических уравнений является *нелинейной.* Для ее решения используется прием линеаризации и метод секущих.

Когда для текущего $R_{l,m}(x)$ п.ч.и. не сходится, т.е. число NN модуль-максимумов экстремумов, чередующих знак и равных по величине, меньше $l+m+2$; или не сходится метод секущих, то начинает работать н.а. с наилучшего аппроксиманта $R_{l-1,m-1}(x)$, последовательно повышая его степени до степеней текущего аппроксиманта до тех пор пока не будет получено не менее $l+m+2$ точек, в которых экстремумы уклонения чередуют знак. После этого снова начинает работать метод п.ч.и. до получения аппроксиманта, имеющего $l+m+2$ равных по модулю и чередующих знак экстремумов-уклонений. Этот аппроксимант принимается за начальное приближение $R^{(0)}_{l,m}(x)$ метода п.ч.и. для нахождения наилучшего приближения $R_{l,m}(x)$, и работа к.а. продолжается до получения наилучшего аппроксиманта требуемой точности.

Следует добавить, что использование описанных выше алгоритмов полиномиального приближения позволяет находить также наилучшие равномерные приближения некоторыми нелинейными выражениями, например, *экспоненциальными* $a_0\, exp\left(a_1x+\ldots+a_nx^n\right)$, *логарифмическими* $ln\left[a_n\left(x^n+a_{n-1}x^{n-1}+\ldots+a_0\right)\right]$ и др., для которых доказаны соответствующие обменные теоремы и получены формулы для пересчета параметров приближения [10].

Для построения наилучшего равномерного приближения функции многих (k) переменных $f(X)=f(x_1,\ldots,x_k)$, которая задана на множестве N точек $\left\{X^{(1)},\ldots,X^{(N)}\right\}$, обобщенными полиномами $F_n(X)=\sum_{i=1}^{n}a_i\varphi_i(X)$ используется метод, который является аналогом метода п.ч.и. Е.Я. Ремеза решения задачи наилучшего чебышевского приближения путем сведения этой задачи к *задаче линейного программирования* с неотрицательными коэффициентами [5], [11]. В соответствующем алгоритме реализуются *прямая* и *двойственная задачи* линейного программирования. При этом главная задача – двойственная, решается модифицированным симплекс-методом с учетом того, что на практике количество уравнений N значительно больше количества неизвестных n, и таблица «расширенного базиса» размера $(n+2, n+4)$ при *модифицированном симплекс-методе* существенно меньше опорной таблицы $(n+2, N)$ прямого симплекс-метода. Для большего повышения эффективности алгоритма и его оптимизации по точности и скорости симплекс-таблица заменяется *сжатой* (более чем вдвое), но равноценной по поданной информации таблицей, которая уже содержит допустимое базисное решение. Кроме этого, в процессе решения двойственной задачи реализуется полуоптимальный вариант перехода от одного допустимого базисного решения к следующему.

В некоторых случаях, особенно при приближении на больших отрезках, целесообразно использовать *кусочные приближения*. При таком приближении, благодаря разбиению всего отрезка аппроксимации $[\alpha,\beta]$ на сегменты $[t_{i-1},t_i]$ ($i=\overline{1,r}$, $\alpha \le t_0 < t_1 < \ldots < t_r \le \beta$) и приближении отдельно на каждом из них заданной функции f наилучшим равномерным аппроксимантом F^*, желаемая точность приближения может быть обеспечена при меньших значениях степени этого аппроксиманта [10].

На протяжении многих лет автором в Институте кибернетики НАН Украины ведутся работы по алгоритмизации методов Е.Я. Ремеза, оценке всех видов погрешностей этих алгоритмов и по созданию соответствующих программных средств [1], [3-8], [10], [12].

Для разработанных алгоритмов и их программных реализаций автором получены *оценки всех видов погрешностей*, а именно погрешности постановки задачи за счет дискретного представления аппроксимируемой функции, неустранимой и вычислительной погрешностей алгоритмов, а также полной абсолютной погрешности решения этих задач. Это позволило значительно повысить точность результатов вычислений (в некоторых случаях на порядок) [8]. Наряду с общепринятой схемой оценки полной погрешности (п.п.) [13] автором предложена также *вторая схема оценки* п.п., которая использует известные априорные сведения о структурных свойствах функции, в частности о ее принадлежности к определенному классу. При этом получены *неулучшаемые* для некоторых классов функций *как априорные, так и апостериорные мажорантные детерминированные оценки* всех видов погрешностей, расчеты которых включены в вычислительные схемы алгоритмов и программ.

Для определения значений указанных оценок п.п. решения чебышевской задачи автором были получены представляющие самостоятельный интерес вспомогательные *результаты из конструктивной теории функций*, в том числе уточняющие известные оценки венгерских авторов.

С целью существенного повышения эффективности разработанных вычислительных алгоритмов осуществлялись также различные процедуры для их модификации, в том числе реализован подход, основанный на применении сегментной (кусочной) аппроксимации разными классами аппроксимантов.

Главное преимущество разработанных автором алгоритмов заключается в том, что с целью повышения их эффективности проведена *оптимизация по быстродействию и по точности*.

Оптимизация по быстродействию достигается за счет следующих факторов:

– алгоритмы основаны на методе последовательных чебышевских интерполяций (п.ч.и.) Е.Я. Ремеза и относятся к алгоритмам подъема, при работе которых нижняя граница выбранных на каждом шаге (n+2)-х максимальных уклонений от функции аппроксиманта порядка n не уменьшается на последующих шагах, что делает их предпочтительнее по скорости сходимости;

– в алгоритмах также реализуется на практике один из трех возможных, а именно оптимальный вариант последовательной замены (n+2)-точечных наборов, на которых осуществляются шаги ч.и., что обеспечивает квадратичную скорость сходимости всего итерационного процесса и позволяет получать результат за 2 – 3 итерации;

– комбинированный алгоритм реализации дробно-рациональной аппроксимации позволяет обеспечить сходимость п.ч.и. также и для дробно-рационального случая, при этом сохраняются действия факторов оптимизации для полиномиального случая.

Оптимизация по точности осуществляется благодаря:

– включению в вычислительную схему алгоритмов расчета как априорных, так и апостериорных оценок полной погрешности, что позволяет корректировать полученные результаты, а именно в зависимости от требуемой точности менять степень полинома, число точек сетки, параметр критерия останова и обоснованно выбирать вид аппроксиманта;

– просмотру при исследованиях поведения уклонений полинома от функций на каждом шаге итерации при переходе к следующему шагу всех точек сетки, при этом учитываются как верхняя, так и нижняя границы величины наилучшего приближения;

– использованию на каждом шаге чебышевских интерполяций для решения систем линейных алгебраических уравнений метода Краута, для которого проведена оптимизация по точности;

– применению в алгоритме дробно-рациональной аппроксимации на каждом j-м шаге ч.и. для решения нелинейной системы уравнений относительно коэффициентов и величины наилучшего приближения специального подхода, основанного на линеаризации системы и использовании метода секущей;

– использование схемы Н.С. Бахвалова вместо схемы Горнера для уменьшения погрешности округлений при вычислениях в точках значений полиномов степени больше 10;

– применению кусочно-полиномиальной аппроксимации для сжатия больших и сверхбольших массивов числовых данных, что значительно повышает точность приближения и обеспечивает высокие коэффициенты сжатия по сравнению с аппроксимацией без разбиения на сегменты.

В алгоритмах применялись также дополнительные приемы повышения их эффективности:

– решение задач аппроксимации как для дискретно, так и для аналитически заданных функций, при этом дополнительно вводится процедура вычисления значений функции в точках дискретизации;

– проведение предварительной аппроксимации с целью повышения точности значений дискретных данных в случаях, когда известны свойства исходной аппроксимируемой функции;

– обеспечение двух «входов» в алгоритмы, что позволяет находить либо аппроксимант заданной фиксированной степени («вход» по степени), либо такой, который обеспечивает заданную точность приближения («вход» по точности);

– использование различных классов аппроксимирующих выражений, наиболее соответствующих характеру поведения («природе») приближаемой функции;

– построение полиномиального приближения с произвольным весом $w(x) \neq 0$;

– проверка в комбинированном алгоритме решения дробно-рациональной задачи перед началом каждой следующей итерации возможности появления случаев вырождения или почти вырождения, когда не удается получить нужное количество точек экстремального базиса. Эти случаи в алгоритме распознаются до начала вычислений вырожденного аппроксиманта, после чего подключаются средства для их устранения;

– применение для аппроксимации функций многих переменных модифицированного симплекс-метода и некоторых процедур, которые обеспечивают значительное уменьшение количества вычислений и повышение точности результатов.

Выводы

Аппарат аппроксимации на протяжении многих лет использовался в составе прикладного программного обеспечения отечественных ЭВМ и применялся для сжатия массивов данных при решении различных прикладных задач (в том числе для оборонных целей), при расчётах характеристик сложных динамических систем, например, при расчёте прочностных характеристик летательных аппаратов для НИИ им. Туполева, расчетах траекторий движения искусственных спутников, транссект и кривых трансконтинентальных переносов загрязнений воздушной среды, токовых состояний водных систем (водоемов, водотоков, Черного моря) в связи с последствиями Чернобыльской катастрофы и др.

В процессе применения аппарата наилучшей чебышевской аппроксимации была не только подтверждена его высокая эффективность, но и проводилось дальнейшее усовершенствование. Более подробная информация о работах содержится в приводимых ссылках.

В результате численной реализации алгоритмов аппроксимации были разработаны программные комплексы на языках программирования ФОРТРАН, Алгол, Паскаль, а также на С++ для отечественного суперкомпьютера с кластерной архитектурой СКИТ. В состав прикладного программного обеспечения СКИТ включены две библиотеки программ: библиотека чебышевской аппроксимации функций одной и многих переменных и библиотека для вычисления с повышенной точностью значений элементарных и специальных функций.

Аппарат чебышевской аппроксимации в последнее время применялся для сжатия *больших одномерных массивов-векторов* (с возможным количеством значений до 10 млн чисел) с целью получения *небольшого числа параметров* аппроксимантов. В результате расчётов были получены большие значения коэффициентов сжатия (в среднем более двух порядков). Эти работы выполнялись в рамках создания Подсистемы аппроксимации для сжатия больших массивов числовых данных в составе Информационно-аналитической системы «Бюджетный комитет», а также в рамках научно-технического проекта по разработке программно-технических комплексов для решения расчётных задач АНТК «Антонов».

В настоящее время для СКИТ разрабатывается *пакет программ аппроксимации* функций одной и многих переменных разными способами приближения: интерполяционным, среднеквадратичным и чебышевским [12], который войдет в состав его Базового прикладного программного обеспечения. Этот пакет имеет ряд существенных преимуществ по сравнению с известными аналогичными пакетами и специализированными библиотеками, такими, например, как Mathcad, Maple, MATLAB, Mathematica, MATHLIB, NETLIB.

Алгоритмы и соответствующие им программные реализации построения наилучших чебышёвских приближений функций одной и нескольких переменных использовались на протяжении многих лет для аналитической обработки массивов данных в разных направлениях науки и техники, в том числе для оборонной тематики. В последние годы аппарат чебышёвской аппроксимации широко применяется в рамках научно-технического проекта по разработке программно-технических комплексов для решения расчётных задач АНТК «Антонов», а также для сжатия больших и сверхбольших массивов-векторов с возможным количеством значений до 10 млн

чисел, что позволяет получать на практике большие значения коэффициентов сжатия (в среднем два порядка). Программные комплексы на основе разработанных алгоритмов включены в состав прикладного программного обеспечения суперкомпьютера СКИТ с кластерной архитектурой [3].

Литература

1. Каленчук-Порханова А.А. Аппарат аппроксимации для анализа и синтеза сложных систем / А.А. Каленчук-Порханова // Пр. Міжнар. конф. «50 років Інституту кібернетики ім. В.М. Глушкова НАН України». – Київ, 2008. – С. 354-361.
2. Ремез Е.Я. Основы численных методов чебышевского приближения / Ремез Е.Я. – К. : Наук. думка, 1969. – 623 с.
3. Каленчук-Порханова А.А. Об одном алгоритме полиномиальной чебышевской аппроксимации / А.А. Каленчук-Порханова // Оптимизация вычислительных методов. – К. : Ин-т кибернетики АН УССР, 1974 . – С. 45-51.
4. Иванов В.В. Об эффективности алгоритмов полиномиальных и дробно-рациональных чебышевских приближений / В.В. Иванов, А.А. Каленчук // Конструктивная теория функций. – София, 1983. – С. 72-77.
5. Каленчук-Порханова А.А. Аппроксимация функций одной и многих переменных / А.А. Каленчук-Порханова // Численные методы для многопроцессорного вычислительного комплекса ЕС. – М. : Изд-во ВВИА им. Н.Е. Жуковского, 1987. – С. 366-395.
6. Каленчук-Порханова А.А. Алгоритмы и анализ погрешности наилучшей чебышевской аппроксимации одной переменной / А.А. Каленчук-Порханова // Теория приближения функций : тр. Междунар. конф. теории приближения функций, (Калуга, 1975). – М., 1977. – С. 213-218.
7. Каленчук-Порханова А.А. Наилучшая чебышевская аппроксимация – алгоритмы и их применение / А.А. Каленчук-Порханова // Сб. трудов Международного симп. «Питання оптимізації обчислень». – Киев, 2009.
8. Каленчук-Порханова А.А. Алгоритмы реализации наилучшей чебышевской аппроксимации – повышение их эффективности // Сб. трудов Международного симп. «Питання оптимізації обчислень» – Киев, 2009.
9. Бахвалов Н.С. Численные методы / Бахвалов Н.С. – М. : Наука, 1973. – 631с.
10. Каленчук-Порханова А.А. Аппарат аппроксимации в составе программного обеспечения суперкомпьютера с кластерной архитектурой / А.А. Каленчук-Порханова, Л.П. Вакал // Искусственный интеллект. – 2009. – № 1. – С. 158-165.
11. Александренко В.Л. Алгоритм построения приближённого равномерно-наилучшего решения системы несовместных линейных уравнений / В.Л. Александренко // Алгоритмы и алгоритмические языки. – М. : ВЦ АН СССР, 1968. – Вып. 3. – С. 57-74.
12. Каленчук-Порханова А.А. Пакет программ аппроксимации функций / А.А. Каленчук-Порханова, Л.П. Вакал // Комп'ютерні засоби, мережі і системи. – 2008. – № 7.
13. Воеводин В.В. Ошибки округления и устойчивость в прямых методах линейной алгебры / В.В. Воеводин. – М. : ВЦ МГУ, 1969.

А.О. Каленчук-Порханова

Найкраща чебишовська апроксимація як ефективний спосіб обробки інформації

Розглядається проблема найкращої чебишовської апроксимації для ефективної обробки інформації. Наводяться обгрунтування переваг алгоритмів апроксимації, що пов'язано з їх оптимізацією за точністю та швидкодією.

A. Kalenchuk-Porkhanova

The Best Chebyshev's Approximation for Effective Treatment of Information

The problem of the best Chebyshev's approximation for effective treatment of information is considered. The advantages of the algorithms elaborated which are connected with their accuracy and performance optimization are presented.

Статья поступила в редакцию 30.07.2009.